Review of Goto et al, 2023

This paper compares results from a single moment cloud microphysics scheme (NSW6 simulations) with those from a double moment scheme (NSW6) in global 6-year 14km resolution simulations. The models are also evaluated using observations and other models.

Both versions perform well against observations of LWP and precipitation in a general sense although a deeper analysis of this evaluation would have been useful – e.g., focussing on certain important areas such as stratocumulus regions that are important for the aerosol-cloud interaction radiative forcing.

It is claimed that NDW6 performs better than NSW6, but I don't think that is justified. It is clear that NSW6 performs better in the tropics and NDW6 at high latitudes. It would be better to be upfront about this and to discuss/examine some of the reasons why this is the case.

Hence, the justification for choosing NDW6 over NSW6 is not that strong from the aerosol results. Perhaps this would be better justified by the likely better representation of cloud microphysical processes in NDW6 (e.g, a better representation of fall speeds, hydrometeor separation, droplet evaporation, etc.?). And the large difference in ERFaci are also a good reason to consider NDW6 - although which one is more accurate is unknown... It might be worth mentioning and considering that the LWP response to aerosol could be constrained/evaluated to some degree with a short simulation of a well-observed volcanic eruption, e.g., like the Holuhraun eruption (Malavelle, Nature, 2017; doi:10.1038/nature22974)

There is a large difference in the aerosol-cloud interaction forcing between the NSW6 and NDW6 simulations, but there is less analysis and discussion devoted to that in the paper. For example, there is no analysis about potential differences in the Twomey effect and cloud lifetime adjustments. E.g., how big might these be in NDW6 vs NSW6? Can you compare the PI to PD CDNC differences for NDW6 to the CCN differences in NSW6? Can you compare changes in the cloud fraction (particularly low altitude clouds) – (this should be possible if the simulations were nudged, otherwise it is likely to be noisy.)

It is highly likely that the difference in the autoconversion scheme between NSW6 and NDW6 is having a large impact on the response of the LWP to aerosols (which is very different between the simulations), but this is not mentioned. It should be mentioned and discussed in the results and summary.

In summary, before publication I think this paper needs a bit of work to add some more relevant analysis and discussion – in particular a fairer discussion of the performance of the single moment scheme vs double moment for aerosols and a more thorough analysis of the differences in aerosol-cloud interaction forcing (ERFaci).

Specific comments

L126 – "although rain does not directly change cloud water in this case" – It's not clear what you mean here or why this part of the sentence is necessary. Are you saying that there is no scavenging/collection of cloud water by rain? Or that it happens through a calculated microphysical process rate rather than some direct assumption? You'll have to elaborate or decide whether this part of the sentence is needed.

L126 – “Thus, the impact of scavenged aerosols on cloud water is inevitably overestimated in single-moment bulk cloud microphysics schemes” – it’s not clear what you mean here either. I can see that there may be an overestimate of the number of droplets present for a given amount of aerosol since droplet numbers cannot evolve and reduce over time in a cloud (partial evaporation, coalescence, etc.) in single moment schemes. What you are saying about scavenging is less clear - are you saying that the impact of aerosol changes is likely to be larger in the single moment scheme due to this direct link between CCN and droplet number? And that this would include a larger effect due to removal of aerosol by scavenging? This would need rewriting to say that if so.

L191 – “sulfate are assumed to have unimodal particle size distributions” – what are the shapes of the size distributions? Lognormal?

L207 onwards – you describe the length of the simulations and some other details here – but this is in the aerosol section, which makes it hard to find. The information on the model run length, spin-up, etc. should be in its own subsection. Also you need some details on whether they were free running or nudged to meteorology?

Plus, the details on the satellite data used for model evaluation should be in its own subsection too and not in the aerosol section.

L209 – “which are obtained from the end of the 1-year aerosol online simulations coupled to NSW6.” – you will need to explain what these “online simulations” are? What do you mean by “online”?

L261 – “However, the seasonal and horizontal biases of the NSW6-simulated LWP are effectively cancelled, and the global and annual mean values of the NSW6-simulated LWP appear closer to the MAC results.” – I can’t see much evidence for the NSW6 LWP values having compensating seasonal and horizontal biases…

L262 – “In terms of the distribution pattern and seasonal cycle, the NDW6-simulated LWP is closer to the MAC result compared to the NSW6 result.” – I don’t see any evidence for this. Although I would be wary of trusting the MAC LWP values too much in precipitating regions like the tropics since precipitation can cause biases.

L275 – “the results in NDW6 are generally better and closer to those of the real atmosphere.” – I’m afraid that I don’t think that you can make this statement with the evidence presented. The differences in precipitation (Fig. 1) between NSW6 and NDW6 are very small and it’s hard to say which is best when looking at the spatial maps for January and July. Similarly for LWP. Perhaps NDW6 is looking a bit better for July for the higher latitudes, but it’s quite hard to tell with the colour range chosen – maybe you need to narrow the colour range on the maps to make the differences clearer? But I think it’s clear from the maps and the zonal means that NSW6 does better in the tropics and NDW6 better at higher latitudes even when taking into account the spatial variability and individual months. This would also need addressing at L517.

L288 – “the recent models participating in the AeroCom Phase-III project” – are these what you term “references” in Figure 4? This should be explained in the caption. It would also be better to label this as “Aerocom models” or similar in the figure(s). Similarly, for Fig. 9 – you need separate markers and labels for Kinne (2019) and Thorsen (2021) and you need to explain why there are several crosses for all-sky and clear-sky for “All” and “Anthropogenic”, but not the other plots in the caption – e.g., are they for different AeroCom model members? Are there also several crosses for Thorsen and what are they if so?

L310 – “the NDW6-simulated RPCW is much closer to the CloudSat-retrieved RPCW” – again, I have to disagree here. It is true for the subtropical regions and higher latitudes, but not for the tropics.

L320 – “The difference in the column burden of SO2 between NDW6 and NSW6 is caused by the chemical loss in the aqueous phase (0.7 TgS yr-1 or +2%) and gas phase (1.1 TgS yr-1 or -7%) and wet deposition (0.4 TgS yr-1 or +24%), as shown in Table A1.” – these values seem different to those quoted in Table A1?

L323 – Need to describe HRM and LRM somewhere in the main paper.

L375 – “example, the IRFari dust values are calculated to be -0.46 Wm-2 (NDW6), -0.57 Wm-2 (NSW6), and -0.24 Wm-2 (Kinne, 2019).” – need to make it clear that these are for all-sky.

L375 – it would be good to add a bit of detail about what kind of data Kinne (2019) and Thorsen (2021) represent. I.e., model aerosol reanalysis, satellite observations, etc.

L376 – “This is partly caused by the weaker absorption of AOT in this study compared to the median value of the AeroCom models” – But it also could be due to the higher dust AOT values?

L379 – “This is inconsistent with the results of the larger column burden and AOT of dust in this study compared to those of the AeroCom models” – you should point out that it is consistent with too little SW absorption, though.

L388 – “For other nonlight-absorbing components, i.e., sea salt and sulfate, the difference in the IRFari values between all-sky and clear sky conditions is very small” – this doesn’t look to be the case? The differences between all-sky and clear-sky are largest for sulphate and sea-salt?

L390 – Make it clear that this is for all-sky. Same for L396.

L451 – “The difference in the ERFaci between NDW6 and NSW6 may be partly explained by a nonlinear relationship of the ERFaci to AOT under the different LWPs, as proposed by Carslaw et al. (2013) who argued that even if the aerosol difference between PI and PD is similar, the value of ERFaci can be larger when the aerosol concentration is lower.”

- Presumably you mean here “for different baseline AOT fields”?
- Also “even if the aerosol difference between PI and PD is similar” would be clearer as “even if the PI to PD aerosol difference for two simulations are similar”.

L455 – “Figure 11 shows that the horizontal distribution of changes in the simulated LWP” – it would be good to introduce the LWP adjustment as another potential factor in causing the difference in the ERFaci between NDW6 and NSW6 (since you have previously suggested that the baseline AOT is a potential cause).

But what about potential differences in the Twomey effect and cloud lifetime adjustments? How big might these be in NDW6 vs NSW6? Can you compare the PI to PD CDNC differences for NDW6 to the CCN differences in NSW6? Can you compare changes in the cloud fraction (particularly low altitude clouds) – this should be possible if the simulations were nudged. Otherwise it is likely to be noisy.

L466 – “Notably, the reason for the differences in the susceptibility between NDW6 and NSW6 should be addressed in future studies.”

- Although it is clear that the difference in the autoconversion scheme between NSW6 and NDW6 is likely to have a large impact here and so this should be mentioned and discussed here and in the summary too.

L496 – “The differences in the dust emissions, dust column burden and SO2, AOT, and IRFari values for total aerosols between NDW6 and NSW6 are larger.” – at L480 you said that the AOT differences were small, so this is a bit of a contradiction.

L508 – “These differences are mainly caused by the difference in the susceptibility of the LWP to AOT” – in the results section this wasn’t stated so clearly. It would be good to mention that in the results too – although do you have evidence that this is the case?

L518 – “and thus, the use of NDW6 is recommended in environmental and climate simulations.”

- The justification regarding the LWP and precipitation performance here is not very strong (see above regarding the NSW6 vs NDW6 performance). Perhaps this would be better justified by the likely better representation of cloud microphysical processes in NDW6 (e.g, a better representation of fall speeds, hydrometeor separation, droplet evaporation, etc.?). And the large difference in ERFaci are also a good reason to consider NDW6 - although which one is more accurate is unknown... It might be worth mentioning and considering that the LWP response to aerosol could be constrained/evaluated to some degree with a short simulation of a well-observed volcanic eruption, e.g., like the Holuhraun eruption (Malavelle, Nature, 2017; doi:10.1038/nature22974)

L529-531 – I can’t see the PCC, etc. values from Goto (2020) listed anywhere – these should be included or quoted in this paper somewhere – ideally in a table so that the reader can compare the new simulations to the old ones.

Table A1 – it would be useful to say what kind of observations the difference references use (satellite, model reanalysis, models, etc.).

L554 – “whereas the difference in the cloud microphysics module does not affect the chemical budget of SO2 oxidation” – presumably you are talking about the cloud microphysical changes within the NSW6 module here rather than the difference between double and single moment (since there are large differences for SO2 between NSW6 and NDW6)? You should make this clearer if so.

Typos / grammar

L18 – “but some aerosol species, especially dust and sulfate, have larger differences among the experiments with NDW6 and NSW6 compared to those among the experiments with different horizontal resolutions, i.e., 14 km and 56 km grid spacing,…”

- Would be better as “but differences between the NDW6 and NSW6 experiments are larger for some aerosol species, especially dust and sulfate, compared to those between experiments with different horizontal resolutions, i.e., 14 km and 56 km grid spacings,…”

L48 – “and hence, an elaboration of both the cloud module and aerosol physics module is required to improve ACI in climate models”. “Elaboration” is perhaps not the best choice of words. I recommend “evaluation” instead.

L52 – “It is promising that convective cloud systems are better represented with finer model resolution when cumulus parameterizations are avoided” – I would recommend this instead: “These results suggest that convective cloud systems are better represented with a finer model resolution for which cumulus parameterizations are avoided”

L72 – “solved” -> “resolved”

L88 – “calculcated” -> “run”

L96 - “incorporated to” -> “incorporated into”

L97 – “were reflected to the version in NICAM.19” -> “was incorporated into the version in NICAM.19”

L107 – “to aerosol physics module” -> “to the aerosol physics module”

L157 – “concentration higher” -> “concentration is higher”

L167 – “dependence of leaf area index” – should be “dependence on leaf area index”?

L175 – “OM” has not been defined.

L176 – “Secondary organic aerosols (SOAs) are assumed to be particles by multiplying the emission fluxes of isoprene and terpenes provided by” -> “Secondary organic aerosols (SOAs) are assumed to form particles, which are calculated by multiplying the emission fluxes of isoprene and terpenes provided by” – also, what are they multiplied by? A constant factor?

L178 – “Parts of SO2 are emitted from volcanic eruptions (Diehl et al., 2012) and are formed from DMS,” – “SO2 is emitted from volcanic eruptions (Diehl et al., 2012) and is also formed from DMS,”

L217 – “These identical datasets were prepared and used in Goto et al. (2020)” – I assume you mean “The same datasets were prepared and used in Goto et al. (2020)”. Otherwise it makes it sound like all of the observational datasets are identical.

L249 – “by satellite results” -> “using satellite data”.

L268 – “CF at the low level" -> "low-altitude CF”

L306 – need to define what WSBC and WIBC are somewhere.

L394 – “which can be caused by its lower lifetime among the references” -> “which may be due to its short lifetime relative to the values from Kinne (2019) (and Thorsen?)”.

L395 – “sea salt is more scavenged by wet deposition” -> “sea salt is scavenged more by wet deposition”

L418 – “but comparable to” -> “but are comparable to”

L422 – “are smaller than”-> “are smaller in magnitude than”

L427 – “fluxes are compared for model evaluations of radiation budget.” -> “fluxes are compared and evaluated.”

L427 – “the global and January averages of the SWCRF” -> “the global averages of the SWCRF for January”. Similarly, for L428 for July and L436 for global averages.

L442 – “by comparing the results between NDW6 and NSW6 under the preindustrial (PI) and the present day (PD)” – better to be clear that this means PI and PD emissions for aerosols and precursor gases. E.g., “by comparing the results between NDW6 and NSW6 for simulations with preindustrial (PI) and the present day (PD) aerosol and precursor gas emissions”. Similarly for the caption of Fig. 11.

L457 – “such as the United States, Europe, and East Asia, increase” -> “such as the United States, Europe, and East Asia, increase in magnitude”

L457 – “the NDW6-estimated ERFaci value is larger negatively than the NSW6-estimated ERFaci” -> “the negative NDW6-estimated ERFaci values are larger in magnitude than the NSW6-estimated ERFaci values”

L460 – “key to understand the difference” -> “key to understanding the difference”

L502 – “whereas those in the sulfate are mainly caused by the wet deposition of SO2.” -> “whereas those in the sulfate are mainly caused by the differences in the wet deposition of SO2.”

L530 – “or” -> “of”

L538 – “Surely,” – wrong choice of phrase here. Would be better as “In support of this,”

Figure E1 – the caption has become jumbled.